Information Visualization for Science & Policy: Engaging Users & Avoiding Bias.

Authors: McInerny, G J[1,2]*, Chen, M[3], Freeman, R[4], Gavaghan, D[1], Meyer, M[5], Rowland, F[6], Spiegelhalter, D. J[7], Stefaner, M[8], Tessarolo, G[9,10], & Hortal, J[10].

*Correspondence to: gmcinerny@hotmail.com

Affiliations:

[1] *Department of Computer Science, University of Oxford, Wolfson Building, Parks Road, OX1 3QD, UK.*

[2] *Computational Science Laboratory, Microsoft Research Ltd, 21 Station Road, Cambridge CB1 2FB, UK.*

[3] *Oxford E-science Research Centre, 7 Keble Road, University of Oxford, Oxford, OX1 3QG UK.*

[4] *Institute of Zoology, Zoological Society of London, Regent's Park, London NW1 4RY, UK*

[5] *Scientific Computing and Imaging Institute, School of Computing, University of Utah, Salt Lake City. Utah, 84112 USA.*

[6] *EMBL, European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SD.*

[7] *Statistical Laboratory, Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge, Cambridge CB3 0WB, UK.*

[8] *Freelance information visualization expert, Lilienthal, Germany. http://moritz.stefaner.eu*

[9] *Departamento de Ecologia, Instituto de Ciências Biológicas – ICB, Universidade Federal de Goiás – UFG, Goiânia, GO, Brazil*

[10] *Dept. de Biogeografía y Cambio Global, Museo Nacional de Ciencias Naturales (CSIC). C/José Gutiérrez Abascal 2, 28006 Madrid, Spain.*

ABSTRACT

Visualisations and graphics are fundamental to studying complex subject matter. However, beyond acknowledging this value, scientists and science-policy programmes rarely consider how visualisations can enable discovery, create engaging and robust reporting, or support online resources. Producing accessible and unbiased visualisations from complicated, uncertain data requires expertise and knowledge from science, policy, computing and design. However, visualisation is rarely found in our training, organisations or collaborations. As new policy programs develop – e.g. 'Intergovernmental Platform on Biodiversity & Ecosystem Services' (IPBES) – we need information visualisation to increasingly permeate both the work of scientists and science-policy. The alternative is increased potential for missed discoveries, miscommunications and at worst, creating a bias towards the research that is easiest to display.

VISUALISATION: EXPLORING AND COMMUNICATING INFROMATION

Visualisations and graphics are the most universally engaging of outputs. Yet the issues of producing informative, engaging and unbiased visualisations (exploratory graphics to publication figures, all the way to interactive web interfaces) have received little attention in the Biodiversity Sciences, or Science-Policy. This is despite huge recent developments in the expertise, knowledge, software, web technologies, and the cultural understanding of both visualisation and data.

These developments come at a critical time. Scientific research and policy are further accelerating investments into understanding, predicting and managing changes in the global environment [1–8]. A crucial information gap has then emerged when scientists and organisations come to explore and communicate the wealth of information being produced [9–11]. Turning vast amounts of often complex data and information (Figure 1) into outputs that scientists can study effectively, and that can then engage diverse users and stakeholders, requires that we value and invest in visualisation and graphics. When subject matter is intangible (e.g. due to scale, complexity or abstraction) [12,13] visualisations play a fundamental role in exploring information and generating understanding [14]. In addition to an open scientific infrastructure [15], visualisation and graphics should be amongst the main priorities for developing modern science and science-policy.

Written science-policy reports are often subject to a "*common approach and calibrated language*" [16]. Such conventions are an essential component of communication strategies and assist with building reputation - for instance, by indicating scientific confidence and framing scenario storylines [16]. The same considerations should apply to visualisations, and actually go further given how easily visuals can engage and influence non-expert audiences across language barriers. Without joined-up strategies for developing and disseminating visualisations and graphics (Box 1), those involved in science and science-policy are missing many opportunities and could bias scientific understanding and policy communications towards that which is easiest to display.

Whether through a lack of training or collaboration, a lack of engagement with visualisation will potentially lead to ineffective and biased visualisations. In an age of heightened scientific scrutiny [17,18], this could impact levels of engagement with science and science-policy, and reduce the reputation of both. To be effective, policy initiatives such as IPBES (Box 1) should ensure investment and innovation in visualisation and visual communications keeps pace with the advances being made in scientific research and science-policy processes. For these reasons, the current poverty of visual communication in science and science-policy deserves a significant response [10]. As put by Fischhoff [19], refusing help in communication deserves heavy criticism because the stakes are so high.

In this article we explore four key issues for increasing the role visualisation plays in science and science-policy, which in turn introduces a host of issues in graphical representation [20], technical implementation [21], multidisciplinary collaboration [22], and user-centred design [23]. Whilst we frame some of our discussion around the newly forming IPBES (Box 1), the arguments and proposals are relevant to the use of visualisations throughout science and science-policy (Box 2). We put forward four suggestions for building capacity in visualisation for our communities (Box 3).

TRUTH AND BEAUTY: WHAT WE HIDE IN VISUALISATIONS

Science can have an awkward relationship with style and beauty. For instance, visualisations that are highly engaging can appear disassociated from data sources [24], appear to advocate particular information by giving it prominence [25], or good visualisations might be interpreted as effort diverted away from the science. However, irrespective of content or function, compelling graphics can also create an impression of truth [26] (a so called "*Cartohypnosis*" [27]) and a lower value or reputation can be attributed to poor designs [28–30]. Any visualisation should be produced with an understanding of these potential biases in audiences' perceptions and take control of them.

Maps - Visualising geo-spatial data is a key example of how an image can both display and hide information. Within maps considerable amounts of content can be attractive and familiar geographic patterns (such as the relative sizing of geographic regions, boundaries, contours, spatial patterns, and other topologies). This potentially distract from the data superimposed upon them (Figure 2). When combined with the processes of analysis and the crafting of maps, it can be difficult to discern what information is being displayed, and the nature of models and data underlying an analysis. For instance, models developed using just a few highly localised data points (Figure 2a) can be extrapolated to far larger regions (Figure 2b) [31], then interpolated to far finer scales than the original data (Figure 2c) and then summarised for geo-political regions (Figure 2d). If we actually take control of the different ways visualisations can influence a user (e.g. differences in design and prominence [32–34], sensual, imaginative and analytical stimuli; see refs in [35],), we can make rigorous design choices that reduce bias and visual rhetoric [36]. For instance, maps might be an obvious means to display geo-spatial information outputs, but not always be the clearest way to explain quantitative features of analysis and its raw outputs.

Reproducible and Reusable resources – In order not to hide information, we must recognise that visualizations are not reality [37]. They are representations of data derived from a suite of transformations, filters and visual encodings that have produced the particular style and storyline of a visualisation. Just like any scientific model, the provenance of these choices should be recorded [38,39] so queries of, and reproducibility from, the source materials [21,40,41] are possible. Any particular visualisation could then be re-used in equivalent comparisons with alternative data sources, or alternative visual encodings can be used with the same data (e.g. map projections [42]).

Uncertainty - Balanced reporting of findings is essential in science and at the science-policy interface [16] but few visualisations convey our ignorance alongside our knowledge [26,43,44]. Omitting uncertainty can promote the apparent precision of data or models, especially if an average or single sample of all possible outcomes is displayed. In science-policy, "*calibrated and traceable*" [16] conventions are used to indicate confidence and uncertainty in text. Many conventions also exist in statistical reporting. However, equivalent guidelines and conventions for visualisations and visualising uncertainty are not currently available. Visualising uncertainty is an active research domain even if it is an unresolved issue in information visualisation research (see below).

DESIGNING FOR NON-SCIENTIFIC AUDIENCES

Science-policy audiences are highly diverse [15,40] and often receive information in far richer digital environments (e.g. online applications, software, games) than science typically provides. The page-limited print layouts of academic journals can impose rigid technical formats onto graphics that limit their re-use [45]. For instance, where huge numbers of individually informative pixels are irretrievably crammed into small rasterised images [46] (Figure 2c) and where graphics are otherwise dependent on text, or a publication's format. Scientific outputs are then produced making numerous assumptions about audiences' numeracy, vocabulary, expertise and level of interest.

Experts and novices can also reason in different ways [20] and might require different design features. Decision and policy makers are obviously a key audience [47] but they too are a highly diverse user group and are not always going to be scientifically or statistically expert [16]. Thus, even if science is freely available (e.g. open, publically available science) it can remain broadly inaccessible because science produces a static explanation of research that often requires specialist expertise to understand. Ideally, science would be able to cater for multiple audiences within interactive devices that allow users to explore scientific knowledge on their own terms.

Interactive visualisations - Richer approaches to communicating scientific information could use visualizations and graphics based on those that enabled scientists' own discovery; for example, by creating exploratory web applications linking scientific data, models and visuals within an interactive tool [21]. Users might then select presentation styles suiting their expertise and knowledge, and select particular abstractions, scales, locations or scenarios based on their own background, interests or serendipitous choices. Such user-driven selections should maintain some connection to the broader context of information. These principles should be applied to all types of information contributed to the IPBES (Box 1). One example comes from the 'Protected Planet' web interface [48] where users are encouraged to edit a community version of data records and rate submitted photographs, when accessing the World Database on Protected Areas (WDPA).

Design Approaches - Scientists rarely come into contact with the full breadth of potential audiences [25] and might not always understand their characteristics and motivations. User-centred [23] and Participatory design approaches [49] explicitly involve stakeholders in the development and design processes, and could better ensure the diversity of user needs are met (Box S1). For instance, policy audiences need to re-communicate information to secondary audiences (e.g. other policy audiences, companies, public & media) and this reuse could be included in the design of visualizations in order to minimise the biases arising through a chain of communicators, especially where scrutiny can increase along that chain [17]. Likewise, ethnographic research and user studies [50] could generate insights that strengthens and shortens the information pathways between stakeholders and that increases the flow of information. Successful design requires realistic consideration of the demands that success may entail [51] - for instance, moving beyond communication of 'facts' towards empowering 'understanding'[52]. Thus, many benefits will come evaluation procedures that reflect diversity in end-users (Box S2). The 'Future Earth' programme is embarking on taking on some of these challenges by developing a "*co-design*" process and by integrating visualisations within any data services provided [3,53]. Given the rarity of this ethos, how the co-design process is developed could be as influential as the end products.

REDUCING THE MULTIDIMENSIONALITY OF COMPLICATED INFORMATION

Most visual interfaces are 2-dimensional (paper, computer displays) and present considerable challenges for displaying complex multidimensional information (Figure 1) [54]. For instance, it can be difficult to include further information (such as uncertainty) into heat maps and choropleth maps (see Figure 2d) because the primary axes are already fixed to the spatial dimensions of the data. Any further information must then be incorporated by elaborating on the map by re-designing the glyphs for each spatial position (see below), or by developing an interactive interface (see above) or using an alternative visualisation design altogether.

Empirical information visualization research has explored some possibilities for displaying complex information [55–57] but there are very many possible design solutions and a single definitive design recipe might not exist (e.g. combinations of colours, glyphs, axes, animations, brushing, layouts, interactions…). Whatever visual strategy is ultimately used, it is important that scientific and statistical details are not altered. For instance, where data is based on multiple models, a summary heat map can be produced from an average 'model ensemble' [58]. However, this design choice can alter the properties of the underlying models through re-scaling (Figure S1) and so introduce a systematic bias into the scientific message.

Interactive exploration and user narratives - Multidimensional information can be difficult for experts to navigate, let alone non-experts. A robust 'mental model' might only develop through a user themselves exploring the complex relationships involved in a system, model, data set or process [25,59]. However, science is strongly biased towards 'explanatory' figures that summarise information, rather than producing ‘exploratory’ knowledge interfaces where audiences can ‘learn by doing’ [60]. One solution for simplifying multidimensional information is to produce a narrative that focuses on a subset of scenarios, data sources, content or otherwise contrasts information in order to create a manageable and informative storyline [61]. The narrative can focus on particular categories in a data set, or particular parameters in a model, in order to reduce guide users’ learning. In principle, users could construct and share narratives themselves through interactive features by selecting components of a data set that interest them [62] (e.g. data filters, or model and scenario selections)(Figure S2, Box 2). For instance, where user interfaces have many options [63], users can select their own visualisation, which could be recorded and then compared to those of other users [64]. Such interactivity should be carefully designed to ensure the resultant narratives, through editing or user interactions, are complementary to the whole scientific message [65–67].

Re-designing components of visualisations – Altering the graphical layouts (e.g. split views, or superimposed and summarised views [68]) and glyphs (data icons and symbols) [69] of a visualisation can offer many effective strategies for reducing the dimensionality of information displays, for instance when communicating any data with estimates of its uncertainty [70]. These design solutions should simplify a visual display, but also maintain an unambiguous relationship between our visual and non-visual terminology (e.g. metrics, definitions, abstractions, uncertainty, ignorance), and the data. Combining multiple information sources into glyphs is one of the most obvious solutions but has many potential issues, such as altering the prominence and interpretation of particular values, producing unwanted clustering and layering effects, or causing the observer to infer unintentional secondary patterns (Figure S1, *i-iv*). Practical design solutions will be broadly applicable rather than restricted to particular data resolutions, or other data characteristics such as spatial pattern. Solutions should also remain simple, such that the graphical cues that users are confronted with are not overloaded and do not render an undecipherable "*visual puzzle*" [71]. Perceptual stress can impede or bias users' comprehension, or at worst cause audiences to disengage. These issues of layout and visual encoding continue to be a hot topic in science and information visualization [22] and visualisation research could be explicitly based on the context of use found in science and science policy.

ADDRESSING A TRANSDISCIPLINARY PROBLEM

Rather than being a design or technical service that can be outsourced as an afterthought, appropriate information visualization and communication strategies must come from early integration of visualisation tools and expertise. For instance, by linking those who contribute to, curate, and analyse data and information sources, to designers, communicators and engineers, and then to those who collate and apply that knowledge (Box 1). Vibrant research programmes exist in each of these domains, but their integration is currently insufficient [22]. If a visualisation and visual communication strategy is to be produced that befits the demands of science-policy programmes such as IPBES (see Box 1) this situation must change. There has not been significant engagement or influence on training within ecology and biodiversity sciences to fill those gaps in expertise [12].

Within visualisation, research programmes do exist in visualizing uncertainty [20,72] and the composition of interactive mapping tools [56]. However, this research often uses different forms of data and uses highly controlled user scenarios that do not necessarily support the challenges that scientists face. In addition, scientists might not actually be aware of this literature. The isolation of these fields then needs to be corrected through an on-going dialogue (e.g. working groups, conferences, collaborations) that can place the requirements of the science and policy into visualisation research, and then use that research. This requires individuals and groups (translators) who can lead the way by verbalising the challenges, translating the research and developing examples that inspire progress.

Enabling multidisciplinary collaborations - To make advances, scientists and science-policy initiatives, (such as the IPBES and Future Earth [3]) must broker collaborations that could produce a joined-up approach to visualisation (Box 1). Potential contributors and collaborators might be unaware of these domains and a clearly defined agenda for engagement that goes beyond stating high level requirements for '*decision support systems*' [73], '*web portals*' [74] and '*user friendly*' resources [18]. We cannot expect visualisation practitioners to passively understand our outputs and practices, nor passively diffuse into key roles in our work. Moreover, science-policy programmes are complex, and might not be well understood. Then, organisations need to work hard to communicate themselves and their goals in ways that are not daunting or hindered by organisational barriers. Plans for resource provision must then account for the eligibility of key contributors (e.g. freelancers, businesses) for funding bids and pose visualisation as more than a service. In sum, a balance must be struck between outsourcing visualisation to experts (which would undoubtedly overlook expertise required from the other domains) and embedding visualisation in all other activities (which would dilute visualisation expertise). We must sow the right seeds if we are to embed the relevant expertise within our scientific and science-policy communities.

Generating impact - It is hard to argue against the huge role visually engaging web interfaces could play in reaching users [75] (Box 2). However targeted user research is needed early on in the process to ensure that the goals are realised. Much can be learned from programmes in '*Open Science*' which aim to increase the accessibility of science [15], but science-policy must also generate significant levels of and user engagement [76]. There are then huge opportunities and large incentives for individuals and organisations to take visualisation seriously. For instance, research can gain increased credibility and influence if it directly addresses stakeholder engagement, and potentially receive increased funding. Both top-down (science-policy; e.g. funding, publishing, hiring, policy development, engagement) and bottom-up responses (scientists; e.g. funding bids, training, collaboration) are needed to improve our visual communications, and the accessibility and usability of research more generally.

CONCLUSIONS AND PRACTICAL STEPS

Success in both science and policy are predicated on reliable and unbiased understanding. Furthermore, our strategies for communicating and curating of knowledge are fundamental to the structure and impact of both science and science-policy interfaces [47,73,77]. Thus, it is highly surprising, if not a major failure, that visualisation and visual communication have been so overlooked in the training of scientists [12] and within the development of science-policy work programmes [10]. Visualization should be supporting the whole information pipeline; from *acquiring and exploring data* and *analysing models*, to the *visual analytics* used to reason across research and assessment activities [13,78], all the way to *storytelling* [61] for communicating background information, results and conclusions. Objective and rigorous visualisations and communications will not be developed without addressing the challenges of their production [12,72].

‘Biological visualization’ offers a great example of successfully embedding visualization into science and science-policy [14,79] – e.g. in producing visualisations that enable exploration of large, complex data sets [80,81] using an explicit understanding of user characteristics when developing visualisations [82], and by offering broader strategies for further progressing the development of biological visualisation [79]. This level of success is enabled by significant levels of visualisation expertise, training, publishing opportunities, and conferences (amongst others), which is not generally the case in our sciences. Like biological visualisation, we should build recognition that visualisation is a highly valued career path in science. So far we have not seized upon the variety of visualisation opportunities available, despite the obvious and immediate benefits that have been available for some time.

Given the topics we have introduced and discussed, we present a number of suggestions to generate some capacity which will allow us to act upon these issues and challenges (Box 3). These suggestions target both top-down and bottom-up responses to the current poverty in information visualisation we see in our sciences. There are many reasons to think progress is possible. For instance, technological and research developments have precipitated significant expertise in information- and data- visualization, information graphics and data journalism. When combined with increased cultural awareness of data, visualization, and informatics (and given the web infrastructure) there are huge opportunities to improve the use of visualisations within and beyond science.

From governments [60] and research organisations to the media [83], communication strategies for complex and uncertain scientific research are being re-considered. These pieces offer the foundations for science and science-policy to build on, and for scientists to work towards. The stage is then set for science and science-policy to become visually astute. What are we going to do about it?

Glossary:

Brushing: Where a user positions the cursor or pointer on a screen to activate a secondary function in an interactive application. For instance, by selecting a subset of data via a mouse which then highlights certain values by changing colour or appearance, or triggering another operation such as activating a label by hovering over a subset of the visualisation.

Choropleth Map: A map visualisation where political regions, biomes, or other areas are colour coded for the value of a variable within those areas (such as a climate variable or population size) (see figure 2d). Unlike a heat map (see below, figure 2b), producing a choropleth map might require further data manipulation to summarise results for the desired boundaries (e.g. averaging or interpolation for those areas) from a gridded model for example.

Co-Design: Defined as "*an active involvement of researchers and stakeholders during the entire research process*" [53]. Within this process, researchers and stakeholders work together when defining research questions, methods and defining a strategy for disseminating results, in order to produce trans-disciplinary and targeted approaches to science-policy [53]. Stakeholders can include academic research, science-policy interfaces, policy makers, funders, governments (regional, national and international), development groups, corporations, businesses and industry, public, and the media [53].

Ethnography: Research seeking to understand individual and cultural responses to tools (e.g. software, new information, methods). Ethnography may investigate how users interpret and understand the tools, build relationships with those tools, as well as define the context of use for these tools in real situations. For instance by understanding how people come in to contact with particular information resources, as well as understanding how they interact with those resources, or share those resources and information. Ethnography is highly complementary to Participatory- and User-Centred-Design methods.

Future Earth: Launched in 2012, Future Earth is an international research programme formed to provide critical knowledge on global environmental change and global sustainability [53].

Glyph: A symbol used to represent information. Simple glyphs could be circles or other shapes used to mark a location in a simple x, y plot. More complex glyphs can encode multiple sources of information by using the different visual channels (shape, size, colour, orientation, brightness, texture, location) in a variety of combinations.

Graphical layout: The relative positioning and sizing of different components of a visualisation. For instance, where multiple graphs or figures are used a layout structures the relationship of the different information sources. The layout may communicate some context, or develop a narrative. Examples include inset graphs, small multiple plots (see below), or linked views in a visualisation.

Heat map: Visualisation using a colour coding system to represent the values of a matrix or grid system (e.g. a gridded map). Heat maps can use a range of colour encodings, or have multiple features where those square glyphs are augmented (see "glyph" above).

Information visualisation: The processes of producing visual representations of data and the outputs of that work. Information Visualisation aims to enhance human's ability to carry out a task by encoding often highly abstract information into a visual form. Visualisations can be static, or interactive and dynamics, and hosted in a variety of media (e.g. journal, poster, website, software).

Intergovernmental Platform for Biodiversity and Ecosystem Services: see Box 1

Linked Views: Interaction where a user interacts with a component of a visualisation that prompts a change in one or more other visualisations. The visualisations can have different axes, glyphs, dimensions or other visual encodings. For instance, one may hover a cursor on a map which feeds that location data to a visualisation highlighting relative rank of that data amongst all locations.

Model Ensemble: A modelled representation comprising of multiple sources of information, more specifically referring to a group of models being used together rather than separately. Each model might be a different method, or use different data sources, or be based on different conditions.

Narrative: A structure developed to reveal information in a particular order, or in particular contrasts, in order to make a point, contextualise information, to pose certain questions, or otherwise create storylines. Narratives can be developed by embellishing graphics and visualisations with annotations, labels or other text, by including other information such as pictures, or via layouts, interactions and animations.

Participatory Design: A process for designing and developing a product that actively involves stakeholders within the whole design process. Unlike 'User-Centred-Design' (see below), participatory approaches can involve greater integration of users in the whole design process.

Science-Policy: The activities and outputs using scientific information to inform and guide general strategies or particular tactics within the policies of governments, NGOs or other organisations.

Small Multiples: A series of graphs using common axes and encodings within a single graphical layout. Small multiples allow different categories to be separated and contrasted where plotting all data simultaneously would result in occluded categories or an otherwise unclear graphic. Small multiples can also be used to develop a narrative (e.g. different patterns evolving through time).

Stakeholder: An individual, group, or organisation that is, or could be, affected by a process or output, or that can affect that process or output. Stakeholders may share a common interest but possibly for very different reasons (such as farmers, agricultural scientists, policy makers).

User-Centred-Design: The process which involves direct interactions with end users when defining, developing and testing a product. From the outset, user requirements are developed so that products are based on a deep understanding of users' education and abilities, as well as their goals, behaviours and motivations, the technology they use, and in what environments (context of use). In contrast to participatory design, users may not be directly involved in the design process.

Uncertainty: Uncertainty can refer to a variety of concepts including ignorance, incompleteness, variation, and stochasticity. Uncertainty can be derived from incomplete knowledge, imperfect methods, sources of measurement or observation bias and propagation of multiple sources of uncertainty.

BOX 1: IPBES - INTERGOVERNMENTAL PLATFORM ON BIODIVERSITY & ECOSYSTEM SERVICES

Following the 2010 UN general assembly, the ‘Intergovernmental Platform on Biodiversity & Ecosystem Services’ (IPBES, www.ipbes.net) has developed around the aims of providing an independent scientific platform for biodiversity, and generating significant policy influence. IPBES will frequently deal with complicated large-scale models and multidimensional data resources [1,78,84] that are challenging for experts to analyse let alone communicate [40,57] (Figure1). Given these goals [84] the IPBES faces some immensely demanding challenges - in addition to providing large-scale scientific assessments, the IPBES must engage diverse audiences with diverse services and outputs, whilst ensuring stakeholder ownership and engagement, and also increasing these activities’ efficiency through effective communication [73]. These are demanding goals. Data visualizations and graphics could enhance all these activities within the policy reports and web interfaces that are intended to make vast amounts of data, assessments and documentation accessible (see main text). By firmly embedding visualisation and graphics into its work programmes, the IPBES can immediately go further than previous science-policy programmes, such as the IPCC (Intergovernmental Panel for Climate Change) [76].

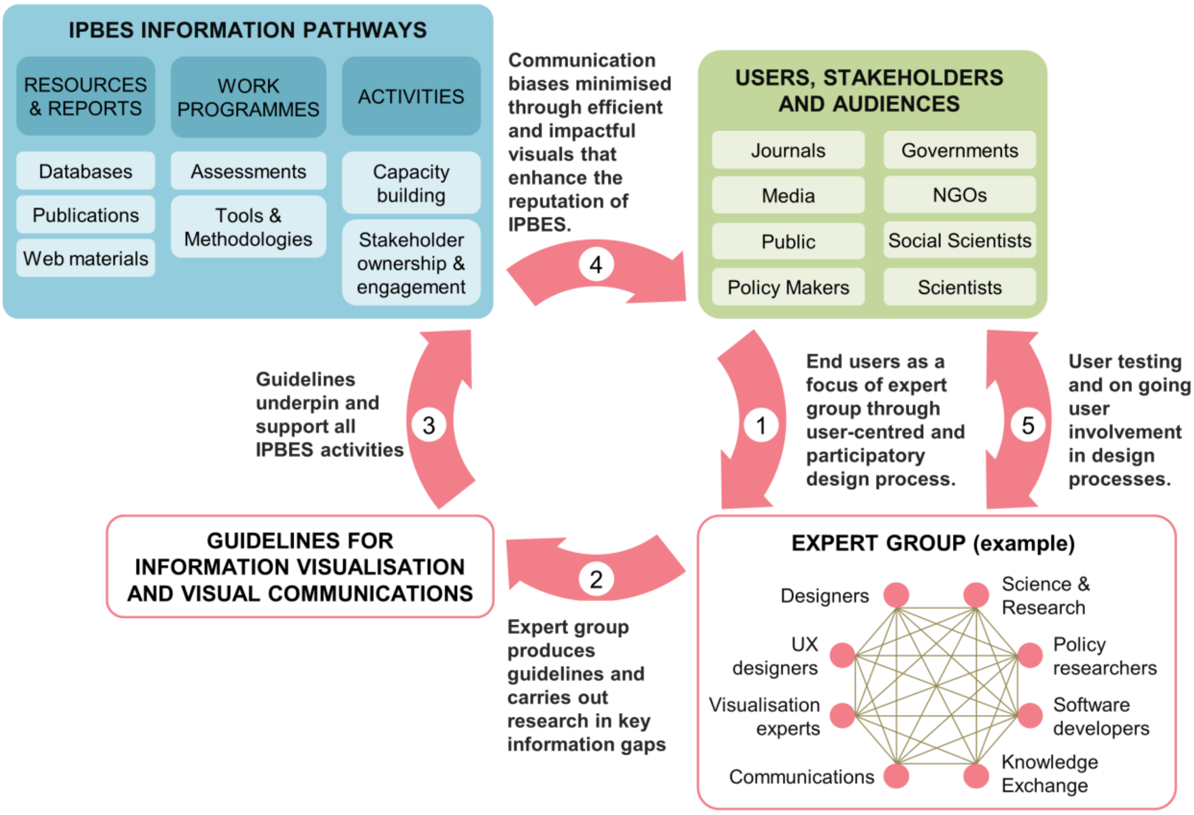


Figure I (w * h; 17.35cm * 11.7cm; 3 column)

An expert group could provide guidelines and strategy which underpin all IPBES outputs and activities. By contextualising communications from the perspectives of end-users, and within the diverse components of the IPBES information pathways, an expert group could help generate efficient and engaging visual communications. As part of a user-centred and iterative design cycle IPBES information pathways could be developed to maximise their effectiveness and impact

---

BOX 2: FROM SCIENTIFIC PAPERS TO INTERACTIVE VISUALISATIONS

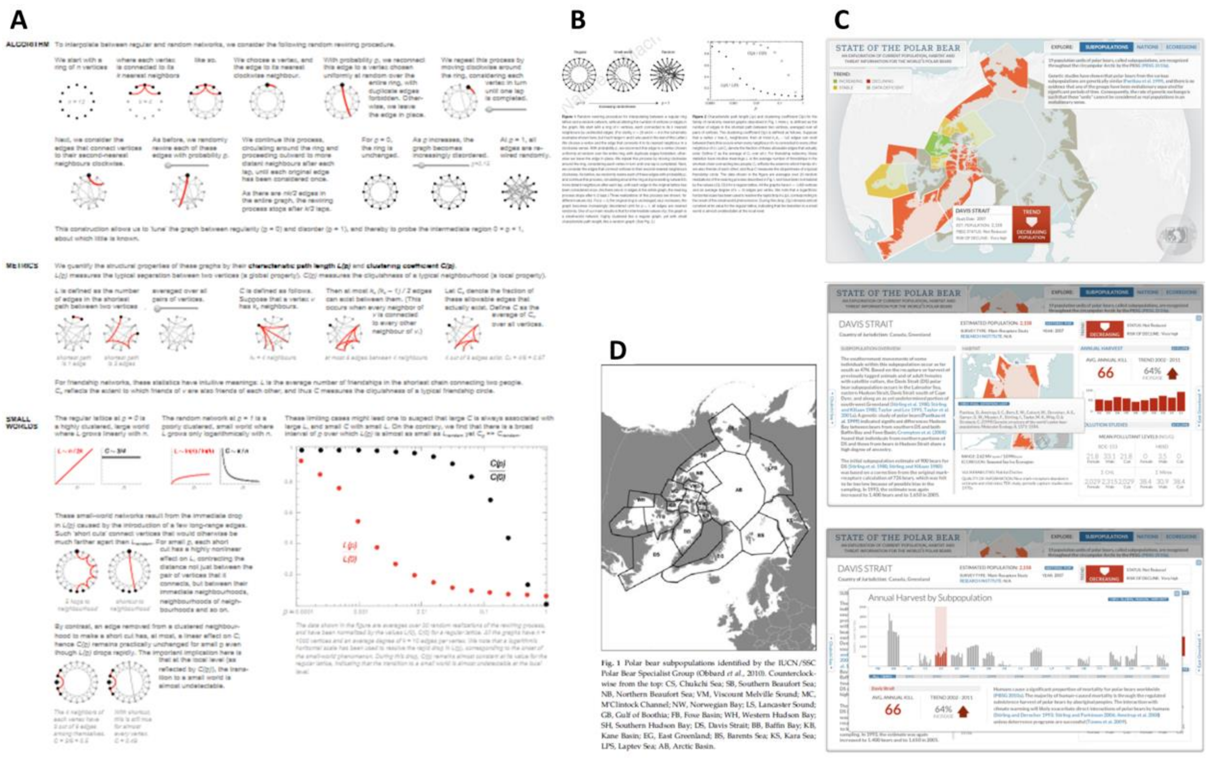


Figure I (w * h; 17.35cm * 10.9cm; 3 column)

Two examples of creating interactive visualisation interfaces alongside graphics from the original scientific papers. In the first example - "*Scientific Communication As Sequential Art*" [85] – (**A**) Bret Victor redesigned a scientific paper, deconstructing the narrative and recomposing it using visualisations, alternative layout and interactive features. Based on the work of (**B**) Watts & Strogatz [86] the 'page' produced by Victor leads a user through the algorithm and metrics that underlie the models being reported. The redesign breaks down the important steps to understanding into manageable steps, which can easily be referred back to as the user develops an overall understanding. Unlike a scientific paper, interactive features allow users to explore the effects of parameters on particular parts of the algorithm or metrics by playing. This example alludes to how a complex theoretical study (or an applied model) could be redesigned using visualisations and dynamic elements to create an accesible through an interactive set of visualisations [17,61,66,67,87].

The second example - "*State of the Polar Bear*" [63] - is an interactive tool designed and developed by the data visualisation company Periscopic, for the IUCN Polar Bear Specialist Group (http://pbsg.npolar.no/en/index.html). The Polar Bear Specialist group advises science-policy and management organisations on the latest scientific knowledge using a variety of information sources that includes more than a thousand articles. (**C**) Within the interactive tool, diverse and fragmented information resources are brought together into a single web application based on interactive visualisations. Users can explore and display data on spatial location, population trends, threats, pollution studies, and harvesting information, and also find refernces upon which this information is based. (**D**) Unlike the scientific literature resources [88], this tool is open access, accessible, dynamic and engaging. In a short time a user can become acquainted with a variety of information sources and through these experiences, and build a picture of the patterns and threats to a species in a way that collections of scientific articles cannot achieve. Also see a new tool - http://globalcarbonatlas.org/ - for exploring carbon fluxes.

BOX 3: FOUR SUGGESTIONS FOR BUILDING VISUALISATION CAPACITY IN OUR COMMUNITIES.

Demand and nurture better quality visualisations and graphics in our science by implementing appropriate training; higher standards for visualisations in journals; and reframing the role of visualisation should play in our scientific work. Increased grass roots expertise will make all other suggestions easier.

Hire expertise and embed it within our organisations in order to seed exemplar projects and work practices; embed expertise that can co-ordinate and deliver appropriate training programs; and to contextualise visualisation research on problems with a direct route to application and further collaboration with visualisation communities.

Embed visualisation in science-policy and knowledge exchange programs by fusing expertise into the processes at an early stage; generating user-requirements and user stories to provide context for the design of visualisations; and producing visualisation and visual communication guidelines that set appropriate standards for designing and evaluating graphics, which should include strategies for engaging further expertise (see below).

Ensure that we can communicate our science and science-policy programmes in appropriate ways to the various areas of expertise that we need to engage – from academic visualisation researchers and visualisation practitioners, to user experience designers and informatics professionals; all the way to designers and communications specialists.

ACKNOWLEDGMENTS

We thank *Microsoft Research Connections* for funding the 2 day workshop "*Visualising the future of our planet - Can we do better than heat maps?*" (organised by GJM, April 2012), and thank all attendees for their contributions during the workshop (Gregor Aisch, David Alderson, Lisa Evans, Mark Freeman, Peter Gassner, Sean Hanna, Jack Harrison, Benjamin Henning, Tim Jupp, Lindsay Lee, Vassily Lyutsarev, Ayman Mogniegh, Rod Page, Mike Pearson, Drew Purves, Tim Regan, Ian Short). We are very grateful to Georgina Mace, Kate Hauxwell, Paul Craze and two anonymous reviewers for comments on the Manuscript. GJM, RF, & DG are part of the "2020 Science" programme (www.2020science.net/) funded by the EPSRC Cross-Discipline Interface Programme (grant number EP/I017909/1) and Microsoft Research.

REFERENCES

1 Clark, J.S. *et al.* (2001) Ecological forecasts: an emerging imperative. *Science* 293, 657–60

2 Perrings, C. *et al.* (2011) Ecology. The biodiversity and ecosystem services science-policy interface. *Science* 331, 1139–40

3 Earth, F. (2013) , Draft initial design report. http://www.icsu.org/future-earth/media-centre/relevant_publications/FutureEarthdraftinitialdesignreport.pdf

4 Scholes, R.J. *et al.* (2008) Ecology. Toward a global biodiversity observing system. *Science* 321, 1044–5

5 Pereira, H.M. *et al.* (2010) Scenarios for global biodiversity in the 21st century. *Science* 330, 1496–501

6 Millenium Ecosystem Assessment. http://www.unep.org/maweb/en/index.aspx

7 http://www.ipcc.ch/

8 Keller, M. *et al.* (2007) A continental strategy for the National Ecological Observatory Network. *Front. Ecol. Envrionment* 6, 282–284

9 Lubchenco, J. (1998) Entering the Century of the Environment : A New Social Contract for Science. *Science.* 279, 491–497

10 Carter, T.R. *et al*. (2007) New Assessment Methods and the Characterisation of Future Conditions. Climate Change 2007: Impacts, Adaptation and Vulnerability. In *Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* (M.L.

Parry, O.F. Canziani, J.P. Palutikof, P. J. van der L. and C. E. H., ed), pp. 133–171, Cambridge University Press, Cambridge, UK

11 Smith, B. *et al.* (2013) COMPASS: Navigating the Rules of Scientific Engagement. *PLoS Biol.* 11, e1001552

12 McInerny, G. (2013) Embedding visual communication into scientific practice. *Trends Ecol. Evol.* 28, 13–14

13 Jetz, W. *et al.* (2012) Integrating biodiversity distribution knowledge: toward a global map of life. *Trends Ecol. Evol.* 27, 151–9

14 Evanko, D. (2010) Foreword. *Nat. Methods* 7, 193–193

15 Society, R. (2012) *Science as an open enterprise - The Royal Society Science Policy Centre report 02/12*,

16 Mastrandrea, M.D. *et al.* (2010) Guidance Note for Lead Authors of the IPCC Fifth Assessment Report on Consistent Treatment of Uncertainties Guidance Note for Lead Authors of the IPCC Fifth Assessment Report.

17 Redford, K.H. *et al.* (2012) Conservation stories, conservation science, and the role of the intergovernmental platform on biodiversity and ecosystem services. *Conserv. Biol.* 26, 757–9

18 Sutherland, W.J. (2013) Review by quality not quantity for better policy. *Nature* 503, 167

19 Fischhoff, B. (2011) Applying the science of communication to the communication of science. *Clim. Change* 108, 701–705

20 MacEachren, A.M. *et al.* (2005) Visualizing Geospatial Information Uncertainty: What We Know and What We Need to Know. *Cartogr. Geogr. Inf. Sci.* 32, 139–160

21 Fox, P. and Hendler, J. (2011) Changing the equation on scientific data visualization. *Science* 331, 705–8

22 Laramee, R.S. and Kosara, R. (2007) 1 . Future Challenges and Unsolved Problems (in Human-Centered Visualization). In *Human-Centered Visualization Environments, Lecture Notes in Computer Science Vol. 4417* (Kerren, A. *et al.*, eds), pp. 231–256, Springer

23 Lieberman, H. Paternó, F. Klann, M. & Wulf, V. (2006) End-User Development: an Emerging Paradigm. In *End-user development* (Lieberman, H, Paterno, F, & Wulf, V., ed), Springer. New York.

24 Kosara, R. (2013) InfoVis Is So Much More: A Comment on Gelman and Unwin and an Invitation to Consider the Opportunities. *J. Comput. Graph. Stat.* 22, 1, 29–32

25 Pidgeon, N. and Fischhoff, B. (2011) The role of social and decision sciences in communicating uncertain climate risks. 1, 35–41

26 Boggs, S.W. (1949) An Atlas of Ignorance: A Needed Stimulus to Honest Thinking and Hard Work. *Proc. Am. Philos. Soc.* 93, 253–258

27 Boggs, S.. (1947) Cartohypnosis. *Sci. Mon.* 64, 469–476.

28 Thorndike, E.L. (1920) A constant error in psychological ratings. *J. Appl. Psychol.* 4, 25–29

29 Kurosu, M. & Kashimura, K. (1995) Apparent usability vs. inherent usability:experimental analysis on the determinants of the apparent usability. *CHI '95 Conf. Companion Hum. Factors Comput. Syst.* 292-293

30 Sillence, E. *et al.* (2004) Trust and mistrust of online health sites. *Proc. 2004 Conf. Hum. factors Comput. Syst. - CHI '04* 6, 663–670

31 Hortal, J. *et al.* (2007) Limitations of biodiversity databases: case study on seed-plant diversity in Tenerife, Canary Islands. *Conserv. Biol.* 21, 853–863

32 Cawthon, N. & Moere, A. Vande (2007) The Effect of Aesthetic on the Usability of Data Visualization. *2007 11th Int. Conf. Inf. Vis.* 637-648

33 McCabe, D.P. & Castel, A.D. (2008) Seeing is believing: the effect of brain images on judgments of scientific reasoning. *Cognition* 107, 343–52

34 Fagerlin, A. *et al.* (2005) Reducing the influence of anecdotal reasoning on people's health care decisions: is a picture worth a thousand statistics? *Med. Decis. Making* 25, 398–405

35 Ginsberg, A.D. (2012) Sensation: in search of aesthetic experience in chemical biology. *Curr. Opin. Chem. Biol.* 16, 553–6

36 Harley, J.B. (1989) Historical geography and the cartographic illusion. *J. Hist. Geogr.* 15, 80–91

37 Korzybski, A. (1933) A Non-Aristotelian System and its necessity for rigour in mathematics and physics. In *Science and Sanity: An Introduction to Non-Aristotelian Systems and General Semantics*, (Korzybski, A., ed), pp. 747–761, Lakeville, CT: International Non-Aristotelian Library

38 Stewart, E.C.A. *et al.* (2009) *Cyberinfrastructure Software Sustainability and Reusability : Report from an NSF-funded workshop Report from an NSF-funded workshop held 27 and 28 March 2009*,

39 Tohline, J.E. *et al.* (2007) Provenance for visualizations: reproducibility and beyond. *Comput. Sci. Eng.* 9, 82–89

40 Peters, D.P.C. (2010) Accessible ecology: synthesis of the long, deep, and broad. *Trends Ecol. Evol.* 25, 592–601

41 Sandve, G.K. *et al.* (2013) Ten Simple Rules for Reproducible Computational Research. *PLoS Comput. Biol.* 9, e1003285

42 Aisch, G. Show case of map projections. http://vis4.net/labs/as3proj/

43 Elith, J. *et al.* (2002) Mapping epistemic uncertainties and vague concepts in predictions of species distribution. *Ecol. Modell.* 157, 313–329

44 Rocchini, D. *et al.* (2011) Accounting for uncertainty when mapping species distributions: The need for maps of ignorance. *Prog. Phys. Geogr.* 35, 211–226

45 Rosindell, J. and Harmon, L.J. (2012) OneZoom: A Fractal Explorer for the Tree of Life. *PLoS Biol.* 10, e1001406

46 Hof, C. *et al.* (2011) Additive threats from pathogens, climate and land-use change for global amphibian diversity. *Nature* 480, 1–6

47 IPCC communications strategy. http://www.ipcc.ch/meetings/session35/IAC_CommunicationStrategy.pdf

48 UNEPWCMC Protected Planet. http://www.protectedplanet.net/

49 Simonsen, S. & Robertson, T. (2012) *Routledge Handbook of Participatory Design*, Routledge.

50 Sedlmair, M. *et al.* (2012) Design Study Methodology : Reflections from the Trenches and the Stacks. *IEEE Trans. Visualization and Computer Graphics*, 18(12): 2431-2440, 2012.

51 Lipkus, I. M., & Hollands, J.G. (1999) The visual communication of risk. *JNCI Monogr.* 25, 149–163

52 Garnett, S.T. and Lindenmayer, D.B. (2011) Conservation science must engender hope to succeed. *Trends Ecol. Evol.* 26, 59–60

53 Rockström, J. & Liverman, D. (2013) *Future Earth Initial Design*, Future Earth.

54 Kaye, N.R. *et al.* (2012) Mapping the climate: guidance on appropriate techniques to map climate variables and their uncertainty. *Geosci. Model Dev.* 5, 245–256

55 Sanyal, J. *et al.* (2009) A user study to compare four uncertainty visualization methods for 1D and 2D datasets. *IEEE Trans. Vis. Comput. Graph.* 15, 1209–18

56 Reusser, D.E. *et al.* (2011) Presentation of uncertainties on web platforms for climate change information. *Procedia Environ. Sci.* 7, 80–85

57 Johnson, C.R. and Sanderson, A.R. (2003) A Next Step : Visualizing Errors and Uncertainty. *IEEE Comput. Graph. Appl.* 23, 6–10

58 Araújo, M.B. and New, M. (2007) Ensemble forecasting of species distributions. *Trends Ecol. Evol.* 22, 42–7

59 KILLMATH Ladder of Abstraction. http://worrydream.com/LadderOfAbstraction/

60 Review, B. Blackett Review of High Impact Low Probability Risks – Government office of science. http://www.bis.gov.uk/assets/goscience/docs/b/12-519-blackett-review-high-impact-low-probability-risks

61 Krzywinski, M. & Cairo, A. (2013) Points of view: Storytelling. *Nat. Methods* 10, 687–687

62 Karl, J.W. *et al.* (2013) Geographic searching for ecological studies: a new frontier. *Trends Ecol. Evol.* 28, 383–384

63 http://pbsg.npolar.no/en/dynamic/app/

64 Walker, R. *et al.* (2013) An extensible framework for provenance in human terrain visual analytics. *IEEE Trans. Vis. Comput. Graph.* 19, 2139–48

65 Kosara, R. *et al.* (2002) Useful Properties of Semantic Depth of Field for Better F + C Visualization. *VISSYM '02 Proceedings of the symposium on Data Visualisation 2002*, 205-210

66 Katz, Y. (2013) Against storytelling of scientific results. *Nat. Methods* 10, 1045–1045

67 Krzywinski, M. and Cairo, A. (2013) Reply to: "Against storytelling of scientific results."*Nat. Methods* 10, 1046–1046

68 Wong, B. (2010) Salience. *Nat. Methods.* 7, 773

69 Maguire, E. *et al.* (2012) Taxonomy-based Glyph Design — with a Case Study on Visualizing Workflows of Biological Experiments. *IEEE Trans. Vis. Comput. Graph.* 18, 2603–2612

70 MacEachren, A.M. *et al.* (2012) Visual Semiotics & Uncertainty Visualization: An Empirical Study. *IEEE Trans. Vis. Comput. Graph.* 18, 2496–2505

71 Tufte, E.R. (1983) Data-Ink and Graphical Redesign. In *The Visual Display of Quantitative Information* 4pp. 91–106, Graphics Press

72 Spiegelhalter, D. *et al.* (2011) Visualizing uncertainty about the future. *Science* 333, 1393–400

73 IPBES, Plenary meeting to determine modalities and institutional arrangements for an intergovernmental science-policy platform on biodiversity and ecosystem services. Second session, P.C. 16–21 A. 2012. http://www.cites.org/eng/com/AC-PC/AC26-PC20/E-AC26-PC20-05-A.pdf

74 Schimel, D.S. *et al.* (2011) *NEON 2011 SCIENCE STRATEGY - Enabling Continental-scale Ecological Forecasting*,

75 UNEP (2012) *Issues for the 21st Century: Results of the UNEP Foresight Process on Emerging Environmental Issues.*,

76 Turnhout, E. *et al.* (2012) Conservation policy: Listen to the voices of experience. *Nature* 488, 454–455

77 Vohland, K. *et al.* (2011) How to ensure a credible and efficient IPBES? *Environ. Sci. Policy* 14, 1188–1194

78 Michener, W.K. and Jones, M.B. (2012) Ecoinformatics: supporting ecology as a data-intensive science. *Trends Ecol. Evol.* 27, 85–93

79 Donoghue, S.I.O. *et al.* (2010) Visualizing biological data - now and in the future. *Nat. Publ. Gr.* 7, S2–S4

80 Meyer, M. *et al.* (2009) MizBee: a multiscale synteny browser. *IEEE Trans. Vis. Comput. Graph.* 15, 897–904

81 Meyer, M. *et al.* (2010) Pathline: A Tool For Comparative Functional Genomics. *Comput. Graph. Forum* 29, 1043–1052

82 Pavelin, K. *et al.* (2012) Bioinformatics meets user-centred design: a perspective. *PLoS Comput. Biol.* 8, e1002554

83 BBC BBC Trust review of impartiality and accuracy of the BBC's coverage of science. http://www.bbc.co.uk/bbctrust/assets/files/pdf/our_work/science_impartiality/science_impartiality.pdf

84 Perrings, C. *et al.* (2011) Ecology. The biodiversity and ecosystem services science-policy interface. *Science* 331, 1139–40

85 http://worrydream.com/ScientificCommunicationAsSequentialArt/

86 Watts, D.J. and Strogatz, S.H. (1998) Collective dynamics of "small-world" networks. *Nature* 393, 440–2

87 Leslie, H.M. *et al.* (2013) How good science and stories can go hand-in-hand. *Conserv. Biol.* 27, 1126–9

88 Stirling, I. and Derocher, A.E. (2012) Effects of climate warming on polar bears: a review of the evidence. *Glob. Chang. Biol.* 18, 2694–2706

FIGURES

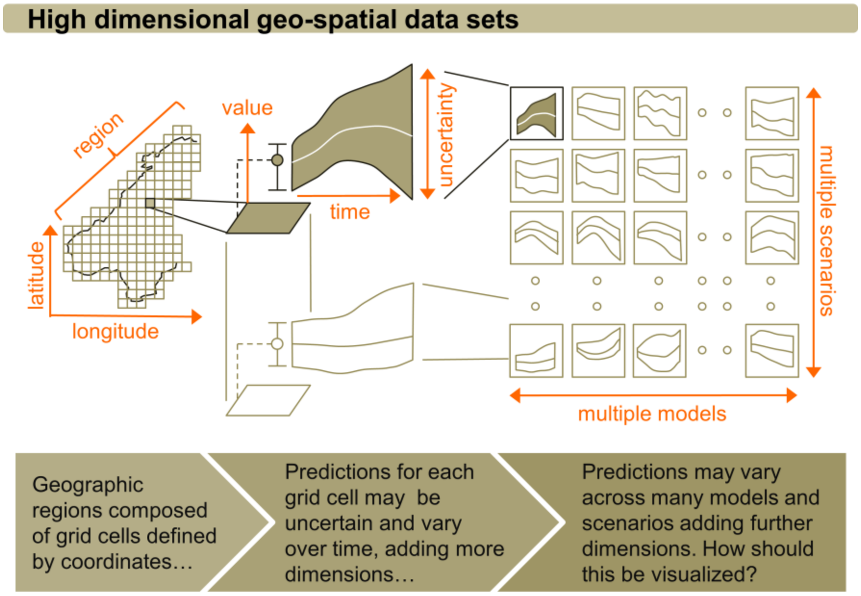


Figure 1 (w * h; 11.3cm * 7.9cm; 2 column)

Scientists and science-policy frequently deal with high dimensional modelled outputs but how will they be visualised? For instance, across spatial regions (e.g. defined by grid cells and spatial co-ordinates) models can predict a value for a metric of interest and which has an associated uncertainty measure, both of which can change over time. When multiplied across a multiplicity of models and scenarios (and also alternative methods and simulations) information displays become highly challenging, even before including meta data or multiple variables of interest and their associated uncertainties.

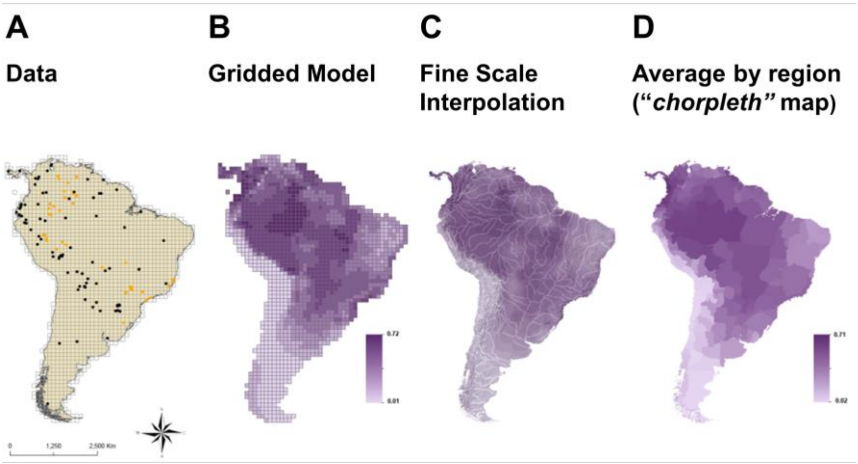
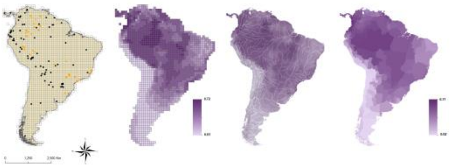


Figure 2 (w * h; 11.3cm * 5.5cm; 2 column)

Highly crafted maps can alter our perception of, and ability to query, models and data. For example, sparse, spatially biased data on a species distribution **(a)** (yellow dots) can be used to create a coarse gridded model **(b)**, which can then projected onto a fine-scale map **(c)** or averaged for geo-political regions **(d)**. Each map confers a different message on the precision and uncertainty of biodiversity information [black dots in **(a)** represent observations not used to develop this model]. Appropriate visual communication techniques must engage users and inform, but also maintain links with the underlying models and data. *See supplementary information for details on this species distribution model of the Jaguar.*

Supplementary Information

Information for Jaguar model shown in figure 2 of main text.



We compare the results of a Species Distribution Model (SDM) based on a biased dataset with an independent source of data, to show that despite the beauty of the maps, they can provide information of poor quality. The geographic distribution of the jaguar (*Panthera onca*) was modeled using all the records of the species available in speciesLink (http://splink.cria.org.b0072/; yellow dots in Figure 2a), a database restricted to Brazil. These data were used to calibrate a SDM with Maximum Entropy Modelling (MaxEnt; Phillips et al., 2006), relating jaguar occurrences at 100 km width grid cells with ten climatic predictors: precipitation of coldest quarter, precipitation of warmest quarter, precipitation seasonality, annual precipitation, mean temperature of wettest quarter, mean temperature of driest quarter, maximum temperature of warmest period, minimum temperature of coldest period, temperature seasonality and annual mean temperature (all obtained from Worldclim; http://www.worldclim.org/, Hijmans et al., 2005).

The climate suitability for jaguar populations predicted by such a model (Figure 2b; the darker the purple tone, the more climatically suitable a given area is) was artificially downscaled to 10 km width pixels, using topographic relief and major rivers to represent major geographic features within the map (Figure 2c). We then compare the geographic distribution of climatic suitability with data from GBIF (http://data.gbif.org/, black dots in Figure 2a), a biodiversity information network that provides occurrence information at a global extent. Note that several occurrences from GBIF are located in areas of low climatic suitability according to SDM results.

R.J. Hijmans *et al*. Very high resolution interpolated climate surfaces for global land areas., *International Journal of Climatology* **25**, 1965–1978 (2005).

S.J. Phillips *et al*. Maximum entropy modeling of species geographic distributions., *Ecological Modelling* **190**, 231–259 (2006).

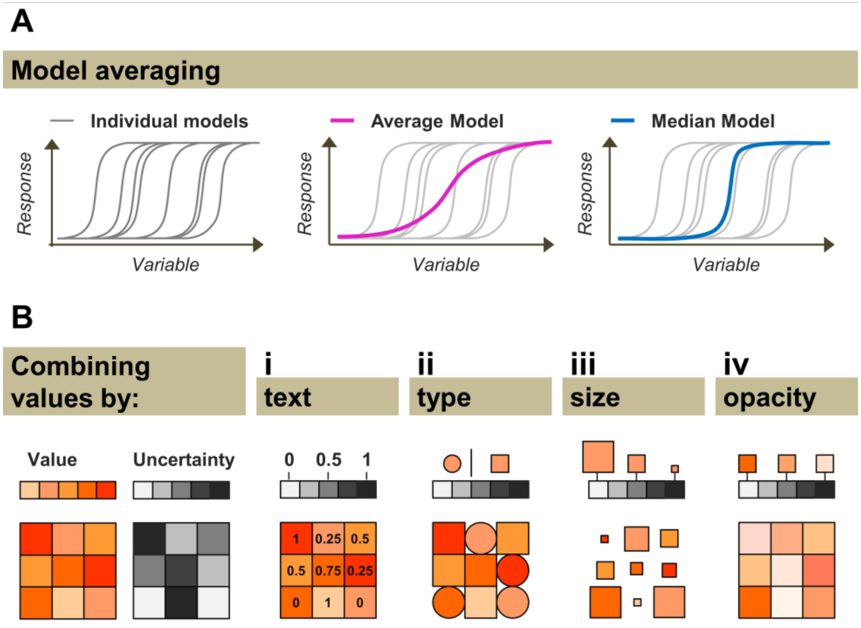


Figure S1

Combining multiple information sources into a single graphic can be challenging. Multiple models may be more easily visualised as an average model, but that average can have different properties from individual models introducing a bias into the communication of the models' properties **(a)**. Other statistics (e.g. median model) can however have similar properties to the individual models and may be more suitable for communication, even if the range of predictions is less well represented. When attempting to integrate 'value' and 'uncertainty' into a single heat map the information may become difficult to read **(b, *i*),** or we can introduce a bias into observers' understanding by causing viewers to perceive layers of values or other secondary patterns **(b, *ii*)**, or altering the prominence of certain values **(b, *iii*)**, or inhibiting observers' ability to assign the meaning of colours to particular a value or level of uncertainty **(b, *iv*)**. Uncertainty is a key focus of policy and visualizing uncertainty is an active, if unresolved, research domain [14]. It may be that the separation of information ('*juxtaposition*') results in the clearest strategy [51], or that having two levels gives the greatest clarity (e.g. **(b *ii*)** high and low uncertainty).

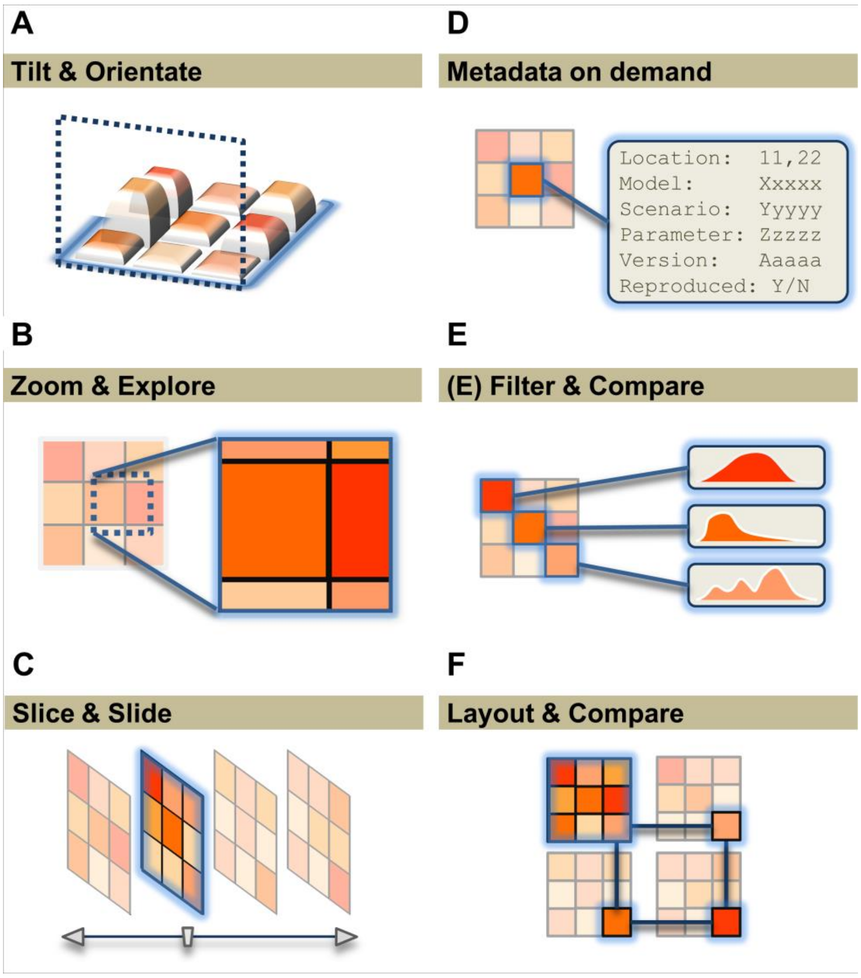


Figure S2

How do we enable users to explore information on their own terms? The ability to interact and create narratives may be vital to engaging users **(a-f)**. These interactions are facets of modern communication applications, such as those alluded to in the IPBES communication strategy [20]. However, interactive displays are not usually supported in the scientific literature, or generated by scientists.

BOX S1: A VARIETY OF DESIGN CHALLENGES

Example audiences:

- ○ Scientist 1 - e.g. domain specialist
- ○ Scientist 2 - e.g. alternative domain
- ○ Politician
- ○ Policy researcher
- ○ Research council
- ○ Lay person 1 - e.g. numerate
- ○ Lay person 2 - e.g. language difference
- ○ Journalist 1 – e.g. scientific
- ○ Journalist 2 – e.g. non-scientific

Example media:

- ○ Printed document
- ○ Scientific publication
- ○ Website
- ○ Poster
- ○ Oral presentation
- ○ Software interface
- ○ TV
- ○ Radio & Internet radio

BOX S2: EXAMPLE MEASURES OF SUCCESS

- ● Audience engagement
- ● Perceptual stress avoidance
- ● Sharing and re-use
- ● Comprehension of information
- ● Developing effective mental models
- ● Reproducibility of information
- ● Comparability with other sources
- ● Citations in science and policy
- ● Views by and impacts on the public
- ● Persistence of recollection and influence
- ● Immunity to developing misleading anecdotes